STA 32 Program 2, due Friday, Jan 31

Problem 1

For this problem, you will use the file `offers.txt`, which contains two columns;

1. Amount (in hundreds of dollars) that the car dealership offered for buying a car.

2. The age group of the person selling the car.

Please name your imported data set "offers". Using this dataset, complete the following problem using `R`:
Create a function which takes in as input the data frame, and a character string called Type which is either "Histogram" or "Boxplot". The function should plot either three histograms, or three boxplots, one for each of the age groups, based on the value of Type.
Please print out the two plots mentioned, and comment on any observed differences in the age groups vs. the amount they were offered for a car.
Hint: The function call should have this basic form (although you have to fill in the body of the if statements):

```
AFunction = function(Data, Type){
if(Type =="Histogram"){

}else if(Type =="Boxplot"){

}else{return("Error: Type not known or unspecified.")}
}
```

Problem 2

Create a function to find the min, mean, variance and mode of a vector (and output the numbers in that order). You may use the functions `min, mean, var` to **check** your answers, but not to calculate the quantities mentioned (you must use `sum` and vector arithmetic to calculate them). Also note that the built in function `mode` doesn't actually give you the mode of the data. Test your function using the first column of the `offers.txt` file. Show your results (for all four), and show they match the results of `R`'s built in functions (for min, mean, and variances).
Hint: To get the mode, you may use the `table, which.max, names` *and* `as.numeric` *functions.*

Problem 3

For this problem, we will use R to simulate the probability of getting a pair (i.e., two of the same card) when you draw two cards from a deck without replacement. Create a function that inputs n, the number of trials, and outputs the probability based on those n trials that a pair was drawn. Calculate the probabilities using $n = 100, 1000, 10000$, and compare that to the true probability (which you have to calculate). Set your seed to 1001 using set.seed(1001) before running your function,i.e.

```
set.seed(1001)
yourfunctionname(10)
yourfunctionname(1000)
yourfunctionname(10000)
```

Problem 4

In this problem we will use R to simulate a biased coin flip. In each trial, you select one of three colored coins (Blue, Red, and Green) with equal probability, and then flip the coin. The Red coin has a 40% chance of landing heads, the Blue coin has a 70% chance, and the Green coin is a fair coin. Create a program which takes in n, the number of trials, and outputs the following probabilities:

Let A be the event that the blue coin was selected. Let B be the event that the coin flip resulted in a head.

(a) $P(A)$

(b) $P(B)$

(c) $P(A \cup B)$

(d) $P(A^C \cup B)$

(e) $P(A \cap B)$

(f) $P(A|B)$

(g) $P(A|B^C)$

Report each probability for n = 100, and 100000. Again set your seed to 1001 before running this function.

Hint: An outline of your program could look like the following:
Step 1: Record the color of the coin selected for n trials.
Step 2: For each of the n trials, based on the color of the coin, assign the appropriate probabilities and flip the coin. (You should now have 2 vectors, one that has recorded color, and one which has recorded head/tail).
Step 3: Use logical arguments to calculate the probabilities above.
Step 4: Output the appropriate probabilities.

How to turn in homework

For problem 1, please print out the plots and attach it to a print out copy of the code for submission.

For problems 2 - 4, I have uploaded a .csv template file named 'program2.csv'. Please fill in your answers for problems 2 to 4 into the corresponding places (I have placed a 0 in the slots where you need to fill in the answers). **Round** your answers to 3 decimal places. Please save the file with the file name "student ID".csv (without quotes of course) (i.e. if your ID is 12345678, then save it as 12345678.csv). Email the file as an attachment. We will use R to check your answers.

Same as the first assignment, copy and paste the code into the body of the email. Make sure you have the appropriate subject line.